Bayesian Clinical Trials for Orthomolecular Medicine

Steve Hickey, PhD;[1] Caroline Chibelushi, PhD[2]

[1] Newlyn Research Group, Newlyn, Cornwall, England; newlynresearchgroup@gmail.com
[2]Staffordshire University, Faculty of Computing, Engineering and Sciences, Staffordshire University, K310, Octagon Building, Beaconside, Stafford, England, ST18 0AD, email: c.chibelushi@staffs.ac.uk

Abstract *Orthomolecular medicine, with its focus on the unique individual, is not well served by large scale trials or statistical analysis of groups. The twin concepts of "rational-patient" and "rational-doctor" as decision makers emphasize the importance of the individual. Rational patients are concerned with efficient and effective interventions to keep them healthy. Similarly, knowledge of the average differences between statistical groups is of limited practical importance to a rational physician. In this paper, we show how Bayesian methods can provide what orthomolecular doctors and patients really need to know: the probability that a treatment will help a patient.*

Introduction

This paper describes how, contrary to conventional medical opinion, small-scale low-cost clinical trials can make medicine more scientific and rigorous. Medicine is currently under the influence of a statistical approach known as *evidence-based medicine* or EBM. The evidence in the name is derived from population statistics, such as EBM's "gold-standard" methods: large-scale clinical trials and meta-analyses.[1] Although these approaches are useful for epidemiology and population health, they do not provide appropriate data for personalized medicine and are easily misapplied to specific patients. By contrast, orthomolecular medicine focuses on biochemical individuality.[2] Thus, its philosophical base diverges fundamentally from the EBM notion of applying averaged statistics to a unique individual.

A rigorous approach to person-centered medicine is possible, using the concept of a *rational-patient*.[3] Rational patients act intelligently, in their own best interests. In other words, they behave reasonably. Similar concepts have been described in other disciplines, such as economics,[4] game theory,[5] and biology.[6] A rational patient wants medicine to provide them with a maximum health benefit, at the lowest cost.

A completely rational patient is not possible, as people are not infinitely intelligent and have a limited amount of time and education to apply to health related decisions. In the real world, people display bounded rationality,[7] and often use simplifying heuristics, or rules of thumb, for practical problem solving.[8] This leads to the concept of a *rational doctor-patient unit*, with the physician acting as an agent or guide, providing specialized information. This approach can ensure each individual patient gets the best available treatment and, indirectly, the patient population statistics are optimized. However, it is important to remember that a treatment that is good on average, and appears best for the group as a whole, may be detrimental to an individual patient; an example of this is a

patient who suffers drug side effects.

Orthomolecular medicine naturally includes the paradigm of the rational doctor-patient unit. Optimal nutrition varies with the individual. A primary issue is predicting what nutritional or related approach is required for one person's needs and health. Such prediction is the core issue in providing effective medicine, but it is not addressed by EBM. Indeed, EBM's statistical approach, based on large groups, is poorly suited to the practical problems of personal medicine.

Fortunately, there are many rigorous methods that can be employed. Here, we illustrate how Bayesian statistics and individual prediction, based on small clinical trials, can be more useful to a rational patient than the so-called high-quality statistical results from EBM.

Prediction or Group Membership?

Medicine covers a broad scope, ranging from the illnesses of individual people through to public health and associated policy development. Those concerned with the health of populations rely on comparisons and descriptions from group statistics. A government agency, for example, might be interested in the average reduction in heart disease following a medical intervention or associated with an element of the diet. For example, in a city with a population of two million, a 0.1% (1 in 1,000) reduction in heart disease might equate to two thousand lives saved. In epidemiology, or randomized clinical trials, large samples provide statistics that are useful for governments or corporate medicine. The task of a family doctor, faced with treating a single patient, is quite different.

Consider Dr. Bob, a hypothetical physician, who wants to know which treatment will help his immediate patient. To be rational, Dr. Bob's decision involves predicting which is the best treatment and estimating the chance that it will benefit his patient. Similarly, Alice, a rational patient, wants a treatment that will help her return to health quickly or will at least offer substantive improvement. Alice considers the value of each proposed treatment heuristically, in a cost-benefit analysis. She expects a treatment for her headache to have a reasonable chance of success. She would not consider a $20 pill, providing a 0.1% (1 in 1,000) probability of relieving her headache, to be a useful or even sensible suggestion.

As a rational patient, Alice would consider the proposal of a practically useless treatment as a reason to change physician. Her acceptable treatment might be an inexpensive pill, offering a 50% (1 in 2) chance of a cure. Her requirements could vary with her condition, as she would consider relief from a migraine more useful than alleviation of a mild tension headache. Alice's working heuristic is that a treatment should provide her at least a 10% benefit. In other words, she would have at least a 1 of 10 chance of substantial benefit, say complete relief from pain.

Sadly, Alice's minimum criterion for benefit is not achieved by current evidence-based medicine, which uses group and population statistics to estimate small effects. Although a treatment with a 0.1% chance of benefit, as mentioned above, might save 2,000 lives in a city, the chance of it helping a particular individual is too small for a rational patient or doctor to choose it. Regardless of its level of statistical significance (which under EBM's methods could be high), Alice is not interested in a treatment that offers her just a 1 in 1,000 chance of being helped.

Contrary to naïve expectations, current evidence-based medicine deals with groups and should not be directly applied to individual treatment. Rational patients want a reasonable prediction that a treatment that will help them *personally*.

Probability

Although most people do not realize it, modern society runs on Bayesian methods. Internet search, electronic communication, and modern finance are just a few of the technologies built around Bayesian techniques. Alan Turing and cryptographers in Bletchley Park also used this approach, which may have shortened the duration of the Second World War by a year.[9] Medicine is one area where Bayesian methods are currently underexploited.

Bayes is of great practical use, because it deals with the "inverse problem." Many clinical studies go from cause to effect, as when a clinical trial tests that a drug (cause) produces a response (effect). An example would be testing to check whether magnesium lowers blood pressure and, if so, by how much. Many important practical and scientific problems work in the opposite direction: we have an effect and want to know the cause. The diagnostic problem is that a patient has symptoms (effects) and the aim is to identify the underlying disease (cause). Bayes addresses such practical questions directly.

To begin our Bayesian approach, we need to reconsider p-values and confidence limits. Our aim is to find the chance that a treatment will benefit an individual patient. Alternatively, we might determine the direct risk of a therapy to a patient. These estimations are surely what a reasonable patient would want to know. Finding the likelihood of a treatment actually helping a patient is more intuitive and direct than the conventional approach. A rational patient like Alice has only an indirect interest in the statistics of groups, she wants to know the probability that the treatment will help her.

In order to provide this information, we need to calculate the chance that Alice's treatment will be helpful and also the chance that it will not. In fact, we can do more than this, we can calculate the likelihood that it will help across the whole range, from no help at all (0% benefit) to total cure (100% benefit). This would us allow us to tell Alice, for example, that there is very little chance the treatment will not help at all, a high chance that it will provide a 60% improvement, and a low to medium chance that it will effect a total cure.

Thus, in predicting the benefits to a patient, Bayes allows us to estimate and assess the probabilities of all outcomes. Every probability value between 0 (no benefit) and 1 (cure) can be considered. For ease of understanding, we report such probabilities as a number out of 100, or a similar fraction. This makes interpretation as a percentage clear, while avoiding the problems with relative values that can mislead the unwary. Direct visual comparison using graphical representations and descriptive statistics are used to aid understanding.

In EBM, results are reported using confidence intervals. The actual value of the parameter, such as a clinical benefit, may lie within the confidence interval, or it may not. Notably, a confidence interval does not actually indicate the probability that the benefit lies between its limits. Counter-intuitively, the benefit parameter is considered to be fixed and its value does not vary.

The standard 95% confidence interval used in EBM means that, if the trial were repeated many times, with sampling from the same population, 19 out of 20 estimated intervals are expected to include the actual value. This implies that in one sample out of every 20, the true value does not fall within the confidence interval. But we only have a single sample. Descriptions of confidence intervals and p-values in EBM often include an explanation of what the intervals are not, since misunderstandings are common. As study size increases, confidence intervals get smaller. This means they can contract around a value that is incorrect because, unlike confidence intervals, systemic errors are not reduced by increasing the number of subjects.

By contrast, Bayesian statistics uses *credible intervals*, which predict that the actual parameter has a particular chance of lying between the limits. The credible interval is thus more intuitively reasonable and agrees with what people expect.[10] With Bayes, the treatment benefit is a variable to be estimated, with its uncertainty indicated by the spread of values. In other words, in a Bayes clinical trial, you can predict the benefit to the patient directly.

For ease of comparison with EBM methods, credible limits can be used in a similar way to the usual confidence limits. Once again, these limits can be expressed naturally; for example, if the chance of a treatment helping a patient is 70%, then 7 in 10 patients will benefit, and if the range is 60%-80%, then 6 to 8 patients out of 10 are expected to benefit. A credible limit thus estimates the range of reasonable values for

the measured benefit.

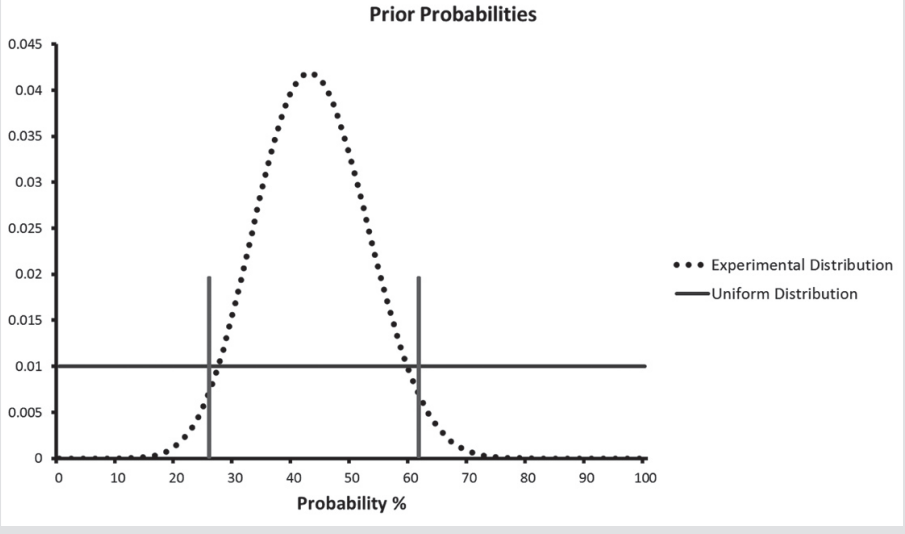
Credible limits, though similar to confidence limits, are easier to understand and can be calculated directly. Using c to indicate the arbitrary "confidence" value (as in c=0.05 or 95% probability), the range of credibility is estimated by calculating the range, 1-c, on both sides of the mean value. **Figure 1,** (below) illustrates calculated 95% credibility limits for a hypothetical experimental distribution, using grey vertical lines.

It is worth noting that Bayesian distributions are not necessarily symmetrical. For consistency with EBM style statistics, credibility limits can be chosen to correspond to a conventional p-value, such as $p<0.05$ (less than 1 in 20), $p<0.01$ (less than 1 in 100), or $p<0.001$ (less than 1 in 1,000). P-values are completely arbitrary, however, and are typically misunderstood. In clinical trials, they are used to indicate a value for the probability of two sample groups being selected randomly from the same parent population. These arbitrary values are used as conventional criteria for significant differences between sample groups. Contrary to the widespread failure of understanding, such significance does not indicate importance or utility, and lack of significance does not indicate the absence of an effect.

Bayes' Theorem

Bayes' approach is relevant to clinical medicine because it mimics the thinking that is engrained in the scientific method. The approach is natural, as it uses new information to update our existing knowledge directly. Computationally, it is more challenging, however, because of the need to calculate the probability distributions and use these for prediction. In recent decades, the increased speed and power of computers has eased this challenge. Currently, EBM's clinical trials are analyzed as stand-alone experiments, with reviews and meta-analyses providing a limited overview. By contrast, Bayesian clinical trials can be considered alone or combined, as each small trial suc-

Figure 1. Bayes can take the pre-experimental belief into account. Here the uniform distribution, which assumes no knowledge or maximum uncertainty, is compared with probabilities determined from later experiments. The vertical grey lines indicate the approximate position of the 95% credibility limits in the experiment.



cessively reduces uncertainty.

Bayes can greatly simplify the statistics of clinical trials. The standard EBM approach requires a large number of arbitrary assumptions, such as a normal or other statistical distribution, confidence limits, p-values, odds ratios, and an array of techniques and tests used in different situations. With Bayes, we apply a single equation to all the data, meaning there is one statistical tool to understand, instead of many. At first Bayes seems a little conceptual and thorny, but this is because people have been taught standard methods. It takes time to become familiar with a different line of attack. However, with time and familiarity, Bayesian methods can appear natural and obvious. A single equation and a little probability can replace all those complicated "evidence-based" medicine statistics.

In considering a clinical trial, we consider the likelihood of the results and our existing knowledge. Using Bayes, we simply multiply the likelihood from the trial by our previous estimate of the probability. This tells us the importance and credibility of the result. So, for example, a recent claim—that high-dose vitamin C increases the risk of kidney stones by a tiny amount (0.147%) in a large population[11]—can be seen as unreliable, considering the substantial prior research indicating the absence of an effect. This contrasts with the current approach, in which each trial report assumes the independent validity of its isolated results, leading to misunderstandings.

Limitations of Bayes

One common idea is that Bayes is subjective and depends on "belief." However, we can use an objective interpretation of Bayes, as in machine learning, where the decisions are essentially made robotically.[12] The term "belief" then indicates a reasonable and rational interpretation of the data, rather than an act of faith.

The other widely reported limitation of Bayesian statistics is that the results depend on the choice of prior beliefs. This objection can be overcome by performing the analysis for a credible range of prior beliefs. This means the effect of our initial assumptions can be determined and quantified. In any event, the initial beliefs issue is overstated, as the updating of the trial with each new patient corrects errors and the credibility increases as additional results are obtained.

Importantly, deciding which of two treatments is more likely to be of benefit to a patient is an undemanding exercise, which does not require great accuracy. For a rational patient, if two treatments for a mild or moderate condition give similar results (say a 50% success rate within 10%, or varying by less than 1 patient in 10), the difference may be deemed to be of little practical importance, as both treatments have a reasonable chance of working. In severe or life threatening illnesses, where a small difference might be more critical, the study can be extended to increase the number of subjects and reduce the uncertainty. However, like David Horrobin,[13] we caution that if a study needs more than one hundred or at most two hundred subjects, the effect we are studying is likely to be too small to be important.

Using Bayesian statistics, the uncertainty in the results falls quickly as the number of patients increases. **Figure 2**, (p.66) illustrates how the distribution of the probabilities of patients with a random treatment (50% effective) narrows, as the number of patients increases from 12 to 96. The peak of the graphs indicates a probability of treatment success of around 50%, as expected.

In an efficient trial, the number of patients should be sufficient to determine a rational or reasonable choice for the clinical problem, but no larger. Clinical research funds are a valuable resource and should be used efficiently and effectively. Large trials are to be avoided. If additional verification is required, another doctor can easily replicate the findings. A study that is easily replicated has greater scientific utility than an overlarge, monolithic trial that wastes resources.[14]

Data and Methods

The simulations and analyses reported here are for illustration only. We programmed the simulations directly in Java, using the inbuilt random number generator, which was sufficiently accurate to produce the required

series of pseudorandom values between 0 and 1. Each pseudorandom value was used to represent a patient's probability, for example, of a treatment being successful. In the resulting Monte-Carlo simulations, the base distributions were objectively estimated for 0.01 intervals, as binomial probabilities. We used Hoffer's vitamin C cancer survival data,[15] and the Hardin Jones analysis by Hoffer and Pauling,[16] to illustrate the use of Bayesian methods in a real world application.

A Sequence of Patients

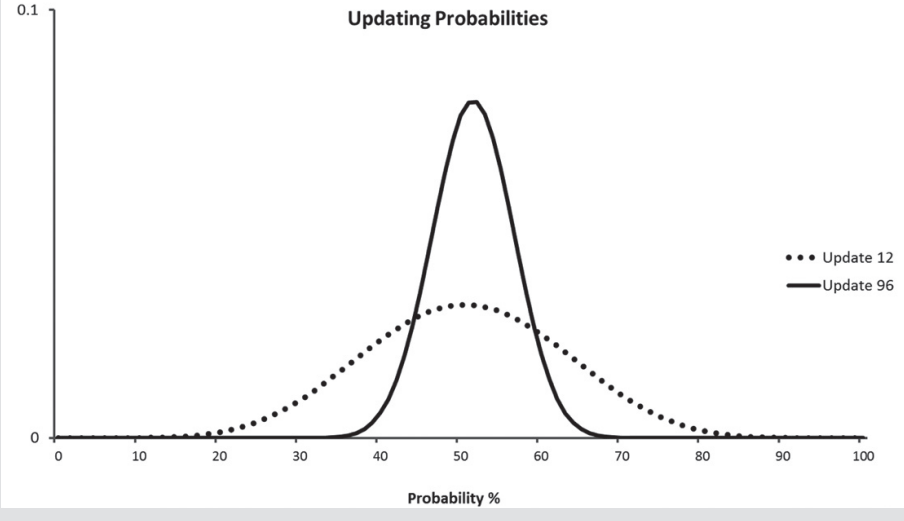
A doctor will typically treat patients in series. Such patients' initial appointments are patient a1, patient a2, patient a3… and so on. Ideally, each patient would be randomly chosen to receive one of two treatments. Where this is not possible, an allocation of sequential patients, so that odd numbers (a1, a3, a5,…) are placed in one group and even numbers (a2, a4, a6,…) in a second, could suffice for choosing group members, as, in practice, the sequence will often be approximately random. For simplicity, we describe the result of treating a sequence of patients as a list of successes, S, or failures, F, as in: S,S,F,S,F,F,F,S,S,…

Note that the method and results apply to a series of single individuals. In EBM, such individual results would be considered "anecdotal," a term which carries connotations of unreliability. Using Bayes, however, we can combine such anecdotal results into a series and our analysis of the outcome can be rigorous.[17] In other words, doctors collaborating from anywhere in the world could combine their case studies to produce a "gold standard" statistical result.

A Simple Trial

Dr. Bob has several patients with a conventionally untreatable and aggressive cancer. About one in four patients ultimately die of the disease. Dr. Bob decides to test two treatments, let us call them Treatment A and Treatment B. Using Bayesian methods, he can do this comparison without a large government grant and with minimal resources. He needs ethical approval for his trial but otherwise the

Figure 2. In a series of hypothetical patients, the width of the distribution reduces rapidly with the number of patients. Bayesian updates for a sequence of 12 and 96 patients are shown.



practical constraints are minimal.

Dr. Bob simply asks his patients if they would like to be involved in the trial and allocates each, randomly, to receive one of the treatments. He is not using placebo controls, as the trial is comparative and thus conforms to the requirements of the Helsinki Declaration of the World Medical Association.[18] Moreover, he is aware that placebos have no effect on definitive outcomes, and death is nothing if not definitive.[19] Dr. Bob decides to use one-year survival as a criterion for his treatments, as he wants a practical indication of efficacy that he can use to help his patients and, in this case, delay is unacceptable. He can follow the patients over longer time scales to update the information, but at the moment he has patients who need help quickly.

Dr. Bob starts by assuming that he has no information about the effectiveness of either treatment. He begins by assuming a complete lack of knowledge and that both treatments are equally likely to succeed or be ineffective. His graph of initial probabilities is an uninformative straight line, as shown in Figure 1. For this simulation, the actual probabilities used were 40% (0.4) for Treatment A and 60% (0.6) for Treatment B. These were chosen to represent a realistic but small advantage to Treatment B.

The graphs are updated as Bayesian probabilities following each patient. The results after the first 12 patients in each group have been followed for a year are shown in **Figure 3,** (p. 68). Visual inspection of the two distributions is informative. From the peaks in the curves, the chance of Treatment A helping a patient is about 43%, whereas the corresponding peak for Treatment B suggests it helps about 50%. While the probability values appear to differ slightly, the large (80%) overlap between the two functions informs Dr. Bob to be cautious in reaching any conclusions.

The reliability of Dr. Bob's estimates increases with the number of patients he treats. Shortly afterwards, Dr. Bob obtained the results in **Figure 4,** (p. 68) with 96 patients in each group. There is now a clear advantage to Treatment B, together with a direct estimate of the chance of it helping the next patient.

At this point, Dr. Bob decides that Treatment B is more likely to be effective. He reports his results to a colleague, Dr. Carol, and continues, using Treatment B for his patients. He has done many Bayesian studies before, and can compare the benefits of them all easily and immediately. He just examines the graphical outputs and checks the corresponding figures. Treatment B is the best treatment he has so far found for helping an individual patient with this aggressive cancer to survive for a year.

Dr. Bob has found a credibly effective treatment, although he would obviously prefer that more of his patients survived. He can continue to assess the treatment or can compare it with alternatives. It is important to understand that the most advanced large-scale randomized placebo controlled trials of "evidence-based" medicine are no more powerful than Dr. Bob's simple test. Indeed, theoretically, even a well-designed EBM trial can merely hope to approximate the utility of Dr. Bob's approach (every admissible statistical procedure is either Bayesian or a limit of Bayesian procedures).[20]

Replicating Dr. Bob

Dr. Carol finds Dr. Bob's results intriguing and wants to use them in her own practice, though she does not believe that Treatment B is superior to the more established Treatment A. Nothing is lost, as Dr. Carol can easily repeat the trial with her own patients. Patients in her new trial will receive one of the two treatments that have the greatest likelihood of helping. Dr. Carol can even extend the trial to include comparison with Treatment C, a promising new therapy that has just become available. Moreover, she can use Dr. Bob's results and those from published research as a starting point (prior probabilities) for her own trial. Importantly, she is not limited to using established results but can compute the effects of "what if" scenarios, such as her initial belief that Treatment A is superior. Thus, she has robust and low-cost methods for checking the credibility of the results.

Figure 3. A simulation of two treatments, each on a sequence of 12 patients, with an actual probability of success of 0.4 (Treatment A) and 0.6 (Treatment B) is shown. There is a large overlap and the experimental distributions are not distinct.

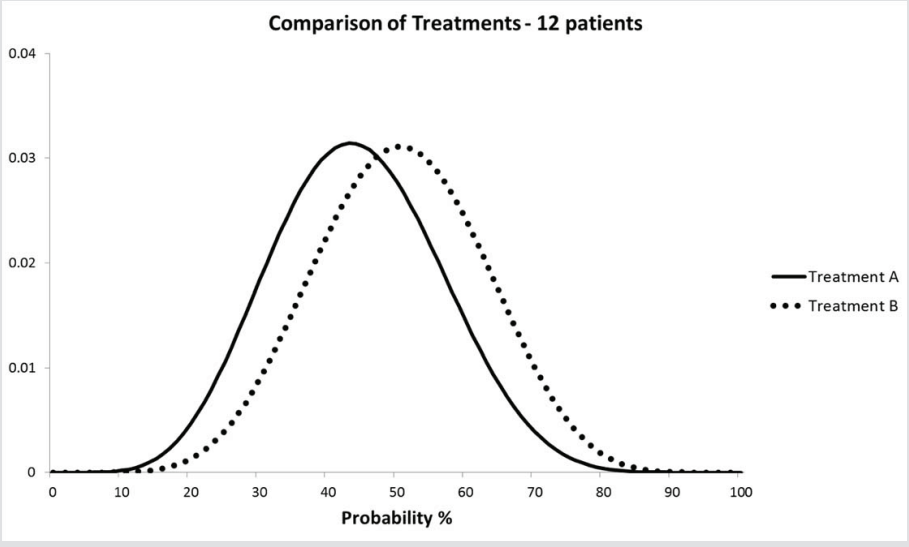
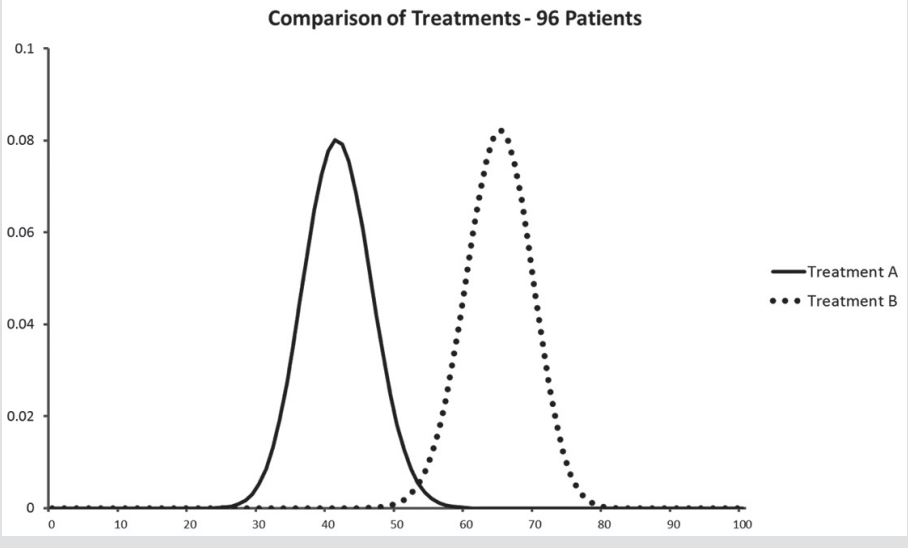


Figure 4. As the number of patients comparing treatments increases the estimates of the probabilities of the treatments helping (peaks in the graph) become more clearly defined. The results for 96 patients indicate a clear advantage for Treatment B over Treatment A. Compare results with those in Figure 3.



Replication and refutation of experiments are the reasons science has been so successful in increasing human knowledge. However, as clinical trials become ever larger and more expensive, the possibility for replication is gradually being eroded and replaced by trust in authorities.

EBM attempts to avoid bias. However, bias can be inadvertently introduced into a large expensive clinical trial or meta-analysis, just as easily as into a small study. The easily replicable small trials we outline can provide high confidence in the utility of medicine, in contrast to expensive, large-scale EBM population-based approaches. Independent replication avoids bias. Bias in one experiment is corrected by the subsequent small trials. Using Bayesian trials, physicians can collaborate worldwide, checking the results and using each additional study to update the medical belief system. There is little or no reliance on governments or large pharmaceutical companies for funds. Importantly, there is less opportunity for fraud or suppression of studies. Any physician, anywhere in the world, has an opportunity to check the findings. Replication and refutation of experiments can once again become part of clinical medicine.

Dr. Hoffer's Cancer Patients

Drs. Pauling and Hoffer provided an interesting Hardin-Jones analysis of the survival of cancer patients, who were given oral vitamin C and some other orthomolecular nutrients. The additional nutrients were somewhat variable (usually vitamin B_3 as niacin or niacinamide at 1.5 or 3 g/day, vitamin B_6 as pyridoxine at 250 mg/day, variable amounts of other B vitamins, vitamin E 800 at IU/day, B-carotene at 30,000 IU/day, selenium at 200-500 mg/day, and other minerals).[16]

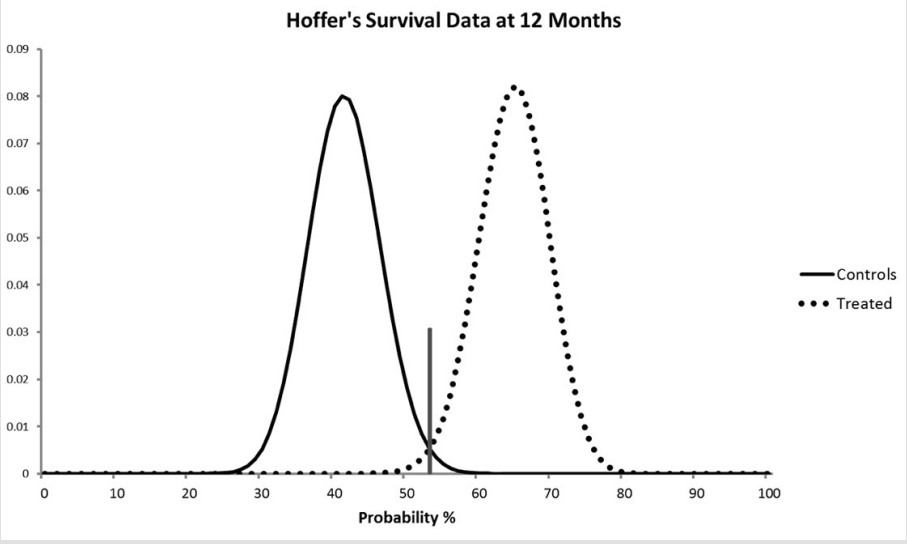
Perhaps surprisingly, the survival of cancer patients can be modeled by assuming a constant risk of death for each individual. Here we show how Bayesian methods can extend the useful information extracted from Hoffer's data. **Figure 5**, (p.70) shows the distribution of 12-month survival probabilities for those treated with vitamin C, and controls. It is immediate from visual inspection that the vitamin treatment increased the chance of survival, as seen by the large separation between the two curves. Displaying such differences visually shows a definite separation of the data and is a basic method for ensuring statistical integrity. Statistical testing can then be used to quantify obvious differences. A working heuristic is that the results are suspect if complicated statistics or large numbers of subjects are needed to determine a result.

Since we have such a clear difference in 12-month survival, we will quantify it. In doing this, we can illustrate the use of decision science[21] and pattern recognition.[22] The vertical grey line in Figure 5 is a decision point for classification and separates the two distributions. The control probabilities (from the control subjects) fall mostly to the left of the line, whereas the treated probabilities (from the subjects receiving vitamin C) fall to the right. The extent to which the distributions overlap, or are on the "wrong" side of the decision line, indicates the uncertainty in deciding that the treated subjects have a greater chance of survival.

The decision line gives a visual representation of the separation. More than 99% of the distribution for survivors taking vitamin C is to the right of the decision line (about 995 or more survivors are expected in 1,000 patients). Correspondingly, less than 1% of the distribution for untreated controls is to the right of the decision line (about 5 or fewer survivors expected in 1,000 patients). The extent of the error (i.e. overlap) is approximately 0.4% by direct numerical integration. Predicting the treatment outcome for individual patients in this way is perhaps the basic requirement for rational physicians and their rational patients.

In **Figure 6**, (p.71) we show the calculated range of Bayesian probabilities for Hoffer's controls, in a 3D map. This is a form of survival curve, considering the range of probabilities at each month of recorded survival. So the survival curve for controls, shown in Figure 5, corresponds to the single horizontal line at 12 months in the surface

Figure 5. Visual comparison of the probability distributions for the survival of Hoffer's cancer patients treated with high-dose oral vitamin C and untreated controls. The treated patients had a large increase in their chance of survival at one year. The vertical grey line is a decision point for prediction (see text).



of Figure 6. For each month, there is a different probability curve and all are displayed in the chart. Note that the curves extend for less than 49 months, at which time there was a single surviving control subject.

Figure 7, (p.71) gives a corresponding plot for those patients Hoffer treated with vitamin C and other nutrients. There are two immediate observations. The probabilities of survival are higher and the computed curves extend out to more than 100 months, at which time several treated patients were surviving, most of whom were reported as "well."

These maps illustrate the potential utility of the Bayesian approach because of the quantity and coverage of the results. Importantly, the results can be recalculated for each quantifiable assumption to estimate the levels of uncertainty or possible bias in our conclusions.

Heuristics for Orthomolecular Trials

Bayesian methods are powerful but are not a substitute for effective design of clinical trials. Donald Berry gave a recent account of Bayesian clinical trials,[23] and there are several introductory books on the background theory,[24] history, and importance of the approach.[9] These references can be used to check how to calculate Bayesian probabilities. The Food and Drug Administration in the U.S. also provide guidelines on using Bayes in clinical trials for medical devices.[25] A large number of related techniques could improve the current EBM approach to medical decision making. These could help introduce reasonable methods of prediction, such as those derived from machine learning.[3,22]

The issue of experimental design remains. Design of effective experiments is an art and is not easily described by a series of fixed rules and methods. Fortunately, heuristics are simple but powerful "rules of thumb" for problem solving. We therefore list some practical heuristics for the design and implementation of effective small clinical trials.

1. Address a clear research question. Before designing the experiment, ensure the question to be answered is clear, unambiguous,

Figure 6. A map of the Bayesian probabilities for Hoffer's controls in his vitamin C studies. The probabilities of survival, as a percentage, are shown for each month. Figure 5 shows the line for a single month (12) in this chart.

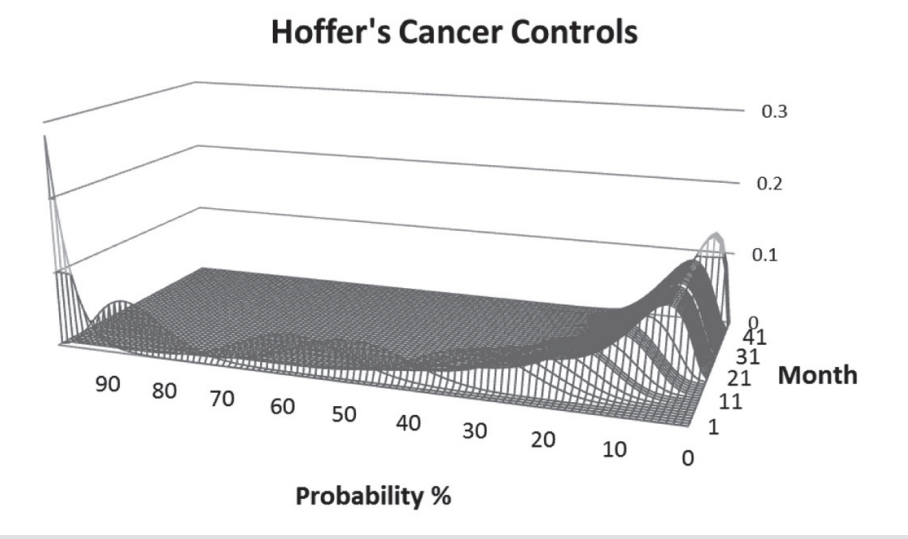
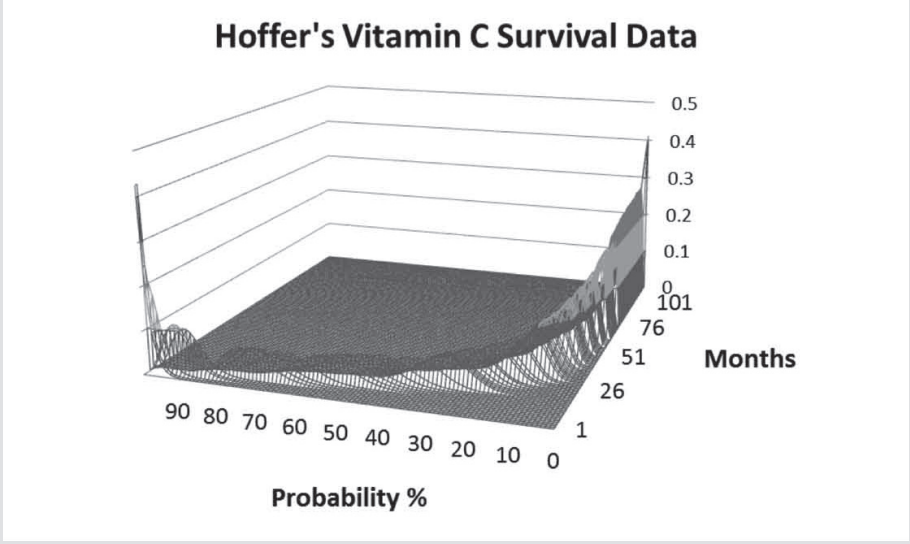


Figure 7. Bayesian probability map of the survival of cancer patients treated orally with high-dose vitamin C, in the Hoffer study. The probabilities of survival, as a percentage, are shown for each month. Compare this map with that of the controls in Figure 6, for a visual representation of the increased chance of survival with treatment.



and if possible decisive. The primary medical question is to predict the treatment that will work for an individual patient.
2. Aim to predict the chance a treatment will help a patient, rather than the probability there is a difference between treated and control groups.
3. Do not consider probabilities after the event, make them part of the study design.
4. A good experimental design determines a *single outcome*. Trials with multiple outcomes are usually weak and often misleading. Each additional outcome needs to be balanced by an adjustment in credible (or confidence) intervals and acceptable probabilities. The useful information provided by the experiment declines above a small number of outcomes.
5. Good experiments measure a *single* or *small number* of parameters. If there are too many parameters, the experimental discrimination falls.
6. A good measurement is direct, for example, blood vessel wall thickness and local blood flow, taken together, measure effective arterial restriction and might be suitable for studying atherosclerosis. Indirect measurements, such as estimates of blood cholesterol, are to be avoided, as they merely provide a vague secondary measure of something that may be correlated with the condition, for some ill-defined reason.
7. The larger and more complicated the trial, the greater is the risk that bias will be introduced.
8. Consider all identifiable sources of possible bias and control for them in the experimental design.
9. Large trials are only needed to detect small benefits; their results are typically immaterial for rational patients wanting practical treatments.
10. Replication is the key to good science. Good experiments are easy to reproduce with a minimum of resources.
11. Comparing large groups of patients gives the average population difference and helps government planners and corporate executives, rather than individual patients.
13. The increased precision in large-scale group comparisons is typically useless for the rational patient or the patient's rational doctor.
14. Testing a sequence of individual patients can be at least as rigorous as comparing groups.
15. Placebos only apply to non-definitive results. Try to compare treatments, rather than using placebos (Declaration of Helsinki).
17. Simple small trials efficiently determine practical differences in treatments and others can easily check the results.
18. Use Bayes, pattern recognition, or any rigorous method of prediction from the decision sciences, and do not be restricted to outdated group statistics by convention.
19. Embrace uncertainty; there is no such thing as "scientific proof" or "best evidence."

Conclusions

Current large-scale clinical trials are expensive, yet are not suitable for personalized medicine. There is thus a need for experiments aimed at measuring and optimizing the benefit to the individual patient. One route to optimization is to make use of all the information available. [3,17] The Bayesian approach makes the efficient use of data much easier. Use of the Bayesian framework, together with small trials, makes it easy to check the correctness of clinical results and evaluate their implications. The use of Bayesian routines in clinical trials is increasingly well received by regulators. At the Biotechnology Industry Organization annual international meeting, FDA Regulatory Counsel indicated that the incorporation of Bayesian reasoning into clinical trial design is likely to be one component of the agency's forthcoming "Critical Path" opportunities list.[26]

Orthomolecular medicine can move forward with a solid scientific base, because of its focus on the individual rather than groups of patients. Governments and large corporations prefer to consider large populations and their management. This leads to an overemphasis on epidemiology and the disproportionate rise of statistical medicine. Nevertheless, key developments in basic sciences are pulling in the opposite direction. In medicine, this pressure for reform is re-

flected in the drive for a person-centered approach, focused on individual genetics and biochemistry, which ultimately leads to orthomolecular ideas.

We can compare this with machine intelligence, where prediction and classification systems are used to make specific decisions in individual cases. The much-vaunted group statistics of EBM were long ago shelved as a method for artificial intelligence, robotics, and much of the decision sciences. Modern prediction techniques will eventually become mainstream: an example is the Watson supercomputer that has gone from winning the game show *Jeopardy!* to medical diagnosis.[27] By embracing a cybernetic approach, small sequential trials and prediction can make orthomolecular medicine a uniquely rigorous or "gold standard" therapeutic discipline.

The rise of EBM may be considered idiosyncratic in the context of the widespread use of advanced prediction techniques in other disciplines.[3] One explanation is the value of EBM to corporate medicine. The absence of direct prediction means that it can be difficult to assess the direct utility of a treatment for a patient. EBM facilitates the continued introduction of marginally useful new drug therapies. Use of indirect group statistics hides the central issue: rational patients want treatments that will help them personally.

As we briefly illustrate, using the Hoffer cancer data, previous findings and published observational data may be analyzed to provide insight and to quantify the current belief system. Bayesian statistics may initially seem complicated to physicians brought up on standard statistics, but they are a natural way to think about the world. Given a little effort, they can be employed to provide increased confidence in findings currently published as anecdotes or uncontrolled observational studies. Orthomolecular clinical trials could benefit from the enhanced rigor provided by the use of Bayesian methods.

Competing Interests

The authors declare that they have no competing interests.

References

1. Straus SE, Glasziou P, Richardson WS, et al: *Evidence-Based Medicine: How to Practice and Teach It.* 4th ed. New York, NY. Churchill Livingstone/Elsevier. 2010.
2. Williams R: Biochemical Individuality: *The Basis for the Genetotrophic Concept.* New Canaan, CT. Keats Publishing. 1998.
3. Hickey S, Roberts H: Tarnished gold. *The Sickness of Evidence-Based Medicine.* Lexington, KY. CreateSpace. 2011.
4. Simon HA: *A behavioral Model of Rational Choice.* Q J Econ, 1955;69:99-118.
5. Davis MD: *Game Theory: A Nontechnical Introduction.* Mineola, NY. Dover Publications. 1997.
6. Smith JM: *Evolution and the Theory of Games.* Cambridge, UK. Cambridge University Press. 1982.
7. Gigerenzer G, Selten R: *Bounded Rationality: the Adaptive Toolbox.* Cambridge, MA. The MIT Press. 2002.
8. Kahneman D: Thinking, fast and slow. New York, NY. Farrar, Straus and Giroux. 2011.
9. Mcgrayne AB: *The Theory That Would Not Die: How Bayes' Rule Cracked the Enigma Code, Hunted Down Russian Submarines, and Emerged Triumphant.* New Haven, CT. Yale University Press. 2012.
10. Bolstad WM: *Introduction to Bayesian Statistics.* Hoboken, NJ. John Wiley & Sons, Inc. 2007.
11. Thomas LD, Elinder CG, Tiselius HG, et al: Ascorbic acid supplements and kidney stone incidence among men: a prospective study. *JAMA Intern Med,* 2013;173:386-388.
12. Barber D: *Bayesian Reasoning and Machine Learning.* Cambridge, UK. Cambridge University Press. 2012.
13. Horrobin DF: Are large clinical trials in rapidly lethal diseases usually unethical? *Lancet,* 2003;361:695-697.
14. Popper K: *The Logic of Scientific Discovery.* New York, NY. Routledge Classics. 2002.
15. Hoffer A, Pauling L: *Vitamin C and Cancer. Discovery, Recovery, Controversy.* Kingston, ON. Quarry Press, Inc. 2000.
16. Hoffer A, Pauling L: Hardin Jones biostatistical analysis of mortality data for cohorts of cancer patients with a large fraction surviving at the termination of the study and a comparison of survival times of cancer patients receiving large regular oral doses of vitamin C and other nutrients with similar patients not receiving those doses. *J Orthomol Med,*1990;5:143-154.
17. Solomonoff RJ: *A Formal theory of inductive inference. Part I. Inf Control,* 1965;7:1-22; Solomonoff RJ: *A Formal theory of inductive inference. Part II. Inf Control,* 1964;7:224-254.
18. World Medical Association Declaration of Helsinki. Ethical Principles for Medical Research Involving Human Subjects. Retrieved from: *[www.*

wma.net/en/30publications/10policies/b3/17c.pdf].
19. Hróbjartsson A, Gøtzsche PC: Placebo interventions for all clinical conditions. Cochrane Database Syst Rev, 2010;(1):CD003974.
20. Wald A: Statistical decision functions. Ann Math Statist, 1949;2:165-205; Wald A: *Statistical Decision Functions.* New York, NY. John Wiley & Sons. 1950.
21. Kleindorfer PR, Kunreuther HG, Schoemaker PJH: *Decision Sciences: An Integrative Perspective.* Cambridge, UK. Cambridge University Press. 1993.
22. Bishop CM: *Pattern Recognition and Machine Learning.* New York, NY. Springer Science+Business Media, LLC. 2006.
23. Berry DA: Bayesian clinical trials. Nat Rev Drug Discov, 2006;5:27-36.
24. Bram U: Thinking statistically. Lexington, KY. CreateSpace. 2012.
25. Guidance for Industry and FDA Staff. Guidance for the Use of Bayesian Statistics in Medical Device Clinical Trials. Retrieved from: [www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/ucm071121.pdf].
26. Tessella Technology & Consulting. Bayesian Statistics in Clinical Trials. The Opportunities and Advantages of Using Bayesian Statistics in Clinical Trials. Retrieved from: [www.tessella.com/wp-content/uploads/2009/11/Capability-statement-Bayesian-Statistics-in-Clinical-Trials.pdf].
27. Giles J: Watson turns medic: supercomputer to diagnose disease. New Sci, 2012;2879. Retrieved from: [www.newscientist.com/article/mg21528796.400-watson-turns-medic-supercomputer-to-diagnose-disease.html].